PATENT

Atty Docket No.: 10008388-1
App. Ser. No.: 10/066,096

Proposed Claim Amendment for Possible RCE

1. (Currently amended) A computerized method for measuring the a degree of coherence of the arrangement of nodes in a hierarchy comprising the steps of:

a) receiving a predetermined hierarchy of nodes arranged in a tree format with one or more subtrees, wherein a subtree of one of the nodes includes any nodes in the hierarchy that stem from the one node, the hierarchy of nodes includes at least a first node and a second node, and the nodes in the hierarchy are associated with one another as one of a sibling node, a child node, and a parent node,

wherein the second node is a child node of the first node when the second node stems from the first node without any intervening node therebetween and the second node belongs in the subtree of the first node;

wherein the first node is a parent node of the second node when the second node is the child node of the first node and belongs in the subtree of the first node; and

wherein the first node is a sibling node with the second node when the first and second nodes stem from a same parent node without any intervening node therebetween, and the first node and the second node belong to the subtree of the same parent node;

b) receiving a plurality of training cases that are filed under said the nodes in the hierarchy; and

c) responsive thereto for determining a (measure of coherence) Coherence Measure for at least one of said nodes in the hierarchy that has a local environment, by evaluating the training cases filed in

the subtree ~~under~~ of the at least one node with respect to the training cases filed in the local environment of the at least one node;

wherein the local environment of the at least one node includes any parent node of the at least one node, any nodes that are sibling nodes of the at least one node, and any nodes that belong to the subtrees under the sibling nodes of the at least one node.

20. (Currently amended) A computerized apparatus for measuring ~~the~~ a degree of coherence of at least one considered node in a hierarchy of nodes that has associated therewith a subtree and a local environment in ~~a~~ the hierarchy comprising:

a) a training case counter for determining the number of training cases under the subtree and the number of training cases for the local environment, the subtree includes any nodes in the hierarchy that stem from the at least one considered node, and the local environment includes any parent node from which the at least one node is stemmed directly, any sibling nodes that are stemmed directly from the parent node of the at least one node, and any nodes that stem from the sibling nodes of the at least one node;

b) an average prevalence determination unit for determining for at least one feature the average prevalence under the subtree and the average prevalence for the local environment; and

c) a predictive feature determination unit for determining the set of predictive features that distinguish training cases of the subtree from documents of the local environment; and

d) a coherence assignment unit for generating a coherence metric number for each considered node based on at least one predictive feature.

22. (Previously Presented) A computerized system for measuring ~~the~~ a degree of coherence of nodes in a topic hierarchy of nodes comprising:

a) a coherence analyzer unit for receiving the topic hierarchy and a set of labeled training cases filed under each of the nodes in the topic hierarcy and responsive thereto for determining, for at least one current node under consideration from the nodes in the topic hierarchy, a measure of coherence in the topic hierarchy of the at least one current node under consideration by evaluating the training cases and at least one feature under ~~the~~ a local environment of the at least one current node and by evaluating the training cases and at least one feature under ~~the~~ a subtree of the at least one current node under consideration;

wherein the subtree of the at least one current node under consideration includes any of the nodes in the topic hierarchy that stem from the at least one current node under consideration; and

wherein the local environment of the at least one current node under consideration includes any of the nodes in the topic hierarchy that stem from a parent node from which the at least one current node under consideration is stemmed directly.

27. (Original) A computerized method for measuring the degree of coherence for ~~one or more current~~ a plurality of nodes in a hierarchy comprising the steps of:

a) receiving ~~a~~ the hierarchy and the training cases filed into said ~~the~~ hierarchy;

b) determining a list of predictive features that distinguish documents of a subtree of a first one of the nodes in the hierarcy ~~the current node's sub-tree~~ from ~~those~~ documents in the current first node's local environment, wherein the first node's subtree includes any nodes in the hierarchy that stem from the first node, and the first node's local environment includes

any parent node from which the first node is stemmed directly, any sibling nodes that are stemmed directly from the first node's parent node, and any nodes that stem from the sibling nodes of the first node;

measure

c) assigning a coherence ~~value~~ to the ~~current~~ first node -based on the list of predictive features and based on one or more of their degree of predictiveness, their ~~the~~ degree of prevalence, and their degree of uniformity, wherein the degree of uniformity reflects how evenly distributed said predictive features are among the ~~children~~ subtrees of the children nodes in the hierarchy that are directly stemmed from the ~~current~~ first node based on the training cases under each of the child subtrees of the children nodes.